Geometric Illustration of the Kalman Filter Gain and Covariance Update Algorithms

Thomas R. Kronhamn

ABSTRACT: The Kalman filter gain and covariance update algorithms are illustrated geometrically for a low dimensional system. The illustrations give an intuitive understanding of the basic filtering process and advocate a more "natural" form of the covariance update algorithm.

Introduction

In his original paper [1], on what later has been commonly called the Kalman filter, Kalman showed the connection between optimal estimates and orthogonal projections. In his derivation of the filter algorithms, the orthogonality of vectors are frequently used. However, neither are geometric figures used as illustrations, nor is the geometric equivalence of the recursive algorithms pursued to full extent. In later estimation literature, tutorials, and textbooks, such as [2] and [3-4] respectively, this is also the case.

However, principles of geometric interpretation of stochastic variables can be found in [3-4]. In [3], a geometric interpretation of the state-estimation problem is also given. Ref. [5] mentions that finding the optimum filter gain is equivalent to minimizing the length of the estimation error vector. However, the geometric interpretation is not used to directly illustrate the Kalman filter gain and covariance update algorithms. For low dimensional, time-discrete systems, this can be done (as shown in this note), thus giving a more intuitive understanding of the basic filtering process. In addition, the geometric representation, as presented, also advocates Eq. (13) as a more "natural" form of the covariance update algorithm than the more commonly used Eq. (12). The form of Eq. (12) is in this sense an algebraically oversimplified form.

Theoretic Preliminaries

The Kalman filter algorithms appear in the literature in many different forms and notations. For convenience, they are briefly repeated here. For more rigorous treatments, readers are referred to the referenced texts, e.g., [5, section 4.2]. Ref. [5] also provides valuable illustrations like the Kalman filter information flow and timing diagrams. Ref. [5] also shares the same basic notation with this paper.

The discrete time, linear system state equation is

\[ x(k) = Ax(k-1) + w(k) \] (1)

where the covariance of Eqs. (5) and (6) are given respectively, by

\[ P(k/k-1) = \text{cov}[\hat{x}(k/k-1), \hat{x}(k/k-1)] = \Phi P(k-1/k-1) \Phi^T + Q(k) \] (10)

\[ D(k) = \text{cov}[\hat{y}(k)], \hat{y}(k) \]

\[ = HP(k/k-1)H^T + R(k) \]

The minimized covariance of Eq. (8) is given by

\[ P(k/k) = \text{cov}[\tilde{x}(k/k), \tilde{x}(k/k)] = [1 - K(k)H]P(k/k-1) \] (12)

We will in the next section, Geometric Illustration, see how Eqs. (9) and (12), or rather another form of Eq. (12), can be found from geometric considerations. This other form of Eq. (12) can also be found by some matrix manipulation of Eq. (12) and with the use of Eqs. (9) and (11). We can then get

\[ P(k/k) = P(k/k-1) - K(k)D(k)K^T(k) \] (13)

This form has several advantages over Eq. (12). One is the relation to the geometric interpretation soon to be demonstrated. Another is that it is symmetric. A sometimes important property for preserving numeric stability. In addition, it is also easy to see that Eq. (13) involves fewer arithmetic operations than Eq. (12) for \( N > M \) (where \( N \) and \( M \) are the number of state and measurement vector elements, respectively).

We now proceed to the geometric representation of stochastic variables. From [3] is found that stochastic variables, e.g., \( x \), and \( y \), can be represented by elements \( x_k \) and \( y_k \) in Euclidean space. The scalar product of \( x_k \) and \( y_k \) is defined by

\[ (x_k, y_k) = \text{cov}(x_k, y_k) \] (14)

The Euclidean norm \( ||x_k|| \) is thus given by

\[ ||x_k|| = (x_k, x_k) = \text{cov}(x_k, x_k) \] (15)

and the angle \( \theta \) between \( x_k \) and \( y_k \) is given

May 1985
This means in short that \( x_s \) is represented by a vector \( x_E \) with length \( \|x_E\| \) defined by Eq. (15) and with an angle \( \theta \) relative to \( y_E \) defined by Eq. (16). From this, it is understood that if \( x_s \) and \( y_s \) are uncorrelated, \( x_E \) and \( y_E \) are orthogonal vectors. The indices \( s \) and \( E \) are omitted in the sequel. The above-defined relations between the stochastic variables and their Euclidean representations should be kept in mind.

**Geometric Illustration**

The Kalman filter gain and covariance update algorithms will now be illustrated in a simple example. Let \( x \) consist of two elements, thus

\[
x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}
\]

With \( H = \begin{bmatrix} 1 & 0 \end{bmatrix} \), \( x_s(k) \) is said to be "directly observable" by \( y(k) \). \( x_s(k) \) is "indirectly observable," meaning that it is estimated only by its correlation to the directly observable element.

**The Directly Observable Element**

Firstly, the residual Eq. (6) is rewritten as

\[
\hat{y}(k) = y(k) - \hat{y}(k/k - 1) = x_s(k) + v(k) - H \hat{x}(k/k - 1)
\]

The measurement noise \( v(k) \) is by definition white and not correlated to the process noise \( w(t) \). The prediction error \( \hat{y}(k/k - 1) \) is thus uncorrelated to \( v(k) \). As seen from Theoretical Preliminaries, this means that \( v(k) \) and \( \hat{y}(k/k - 1) \) are represented geometrically by orthogonal vectors. Equation (18), as a whole, can then be illustrated by the outer triangle in Fig. 1. The state update Eq. (7) can for this state vector element be written

\[
\hat{x}_s(k/k) = \hat{x}_s(k/k - 1) + K_s(k) \hat{y}(k)
\]

and with Eqs. (5) and (8) we get

\[
\hat{x}_s(k/k) = \hat{x}_s(k/k - 1) - K_s(k) \hat{y}(k)
\]

Equation (20) is illustrated in Fig. 1 for the optimal \( K_s(k) \), i.e., that gives minimum

\[
\| \hat{x}_s(k/k) \|.
\]

Minimum \( \| \hat{x}_s(k/k) \| \) is, of course, guaranteed as \( \hat{x}_s(k/k) \) is orthogonal to \( \hat{y}(k) \). This \( K_s \) can now be found from uniform triangles in Fig. 1 as

\[
K_s(k) = \frac{\| \hat{x}_s(k/k - 1) \|}{\| \hat{y}(k) \|}
\]

which gives

\[
K_s(k) = \frac{\| \hat{x}_s(k/k - 1) \|^2}{\| \hat{y}(k) \|^2}
\]

The right-hand side of Eq. (22) may now be transformed to its stochastic counterpart by Eq. (15). This gives

\[
K_s(k) = \frac{\text{cov}[\hat{x}_s(k/k - 1), \hat{y}(k/k - 1)]}{\text{cov}[\hat{y}(k), \hat{y}(k)]}
\]

where \( A_{ij} \) is the \( i,j \)-th element of the matrix \( A \). This means that \( v(k) \) and \( \hat{y}(k) \) are represented geometrically by orthogonal vectors. Equation (18), as a whole, can then be illustrated by the outer triangle in Fig. 1. The state update Eq. (7) can for this state vector element be written

\[
\hat{x}_s(k/k) = \hat{x}_s(k/k - 1) + K_s(k) \hat{y}(k)
\]

and with Eqs. (5) and (8) we get

\[
\hat{x}_s(k/k) = \hat{x}_s(k/k - 1) - K_s(k) \hat{y}(k)
\]

The vector representing the error of the indirectly observable element can be visualized by the two large triangles in Fig. 2. It is defined by its length \( \| \hat{x}_s(k/k - 1) \| \), the angle \( \theta \) to the vector \( \hat{y}(k/k - 1) \) and the fact that \( \hat{x}_s(k/k - 1) \) is orthogonal to \( v(k) \).

Now we introduce the two orthogonal components \( \hat{x}_{s2}(k/k - 1) \) and \( \hat{x}_{s3}(k/k - 1) \) of \( \hat{x}_s(k/k - 1) \) as shown in Fig. 2. This means that

\[
\hat{x}_s(k/k - 1) = \hat{x}_{s2}(k/k - 1) + \hat{x}_{s3}(k/k - 1)
\]

\[
\| \hat{x}_s(k/k - 1) \|^2 = \| \hat{x}_{s2}(k/k - 1) \|^2 + \| \hat{x}_{s3}(k/k - 1) \|^2
\]

\[
\text{cov}[\hat{x}_s(k/k - 1), \hat{x}_s(k/k - 1)] = \text{cov}[\hat{x}_{s2}(k/k - 1), \hat{x}_s(k/k - 1)]
\]

\[
= P_{s2}(k/k - 1)
\]

\[
\text{cov}[\hat{x}_{s3}(k/k - 1), \hat{x}_s(k/k - 1)] = 0
\]

As said at the beginning of Geometric Illustration, an indirectly observable element is estimated by its correlation to a directly observable element. If \( \hat{x}_{s2} \) had been the only

---

**Fig. 1.** Illustration of the measurement residual and the optimal gain of a directly observable state vector element.

**Fig. 2.** Illustration of the measurement residual and the optimal gain of an indirectly observable state vector element.
error component of \( \tilde{x}_2 \) would have been uncorrelated to \( \tilde{x}_1 \) and could not have been estimated.

Now on the other hand, \( \tilde{x}_{2p} \) can be "reached" from \( \tilde{y}(k) \) because it is totally correlated to \( \tilde{x}_1 \). However, this will not affect \( \tilde{z}_0 \).

If we proceed analogously to The Directly Observable Element section, we find the optimal gain from uniform triangles as which gives

\[
K_2(k) = \frac{P_{2,1}(k/k-1)}{D(k)} \quad (34)
\]
as expected from Eq. (9).

That the variance of \( \tilde{x}_2(k/k) \) in this way is minimized can be seen from

\[
\tilde{x}_2(k/k) = \tilde{x}_2(k/k-1) + K_2(k)\tilde{y}(k) \quad (35)
\]
which gives

\[
\tilde{x}_2(k/k) = \tilde{x}_2(k/k-1) - K_2(k)\tilde{y}(k)
= \tilde{x}_2(k/k-1)
+ \left[ \tilde{x}_{2p}(k/k-1) - K_2(k)\tilde{y}(k) \right]
= \tilde{x}_2(k/k-1) + \tilde{x}_{2p}(k/k) \quad (36)
\]
\( \tilde{x}_2(k/k) \) is not explicitly shown in Fig. 2 but can be constructed by vector addition of the two components on the right-hand side. However, \( \tilde{x}_2(k/k) \) is shown to consist of two orthogonal components of which one is not observable from \( \tilde{y}(k) \) and the other is minimized by the actual choice of \( K_2(k) \).

The variance of \( \tilde{x}_2(k/k) \) can be found from Fig. 2 analogously as in The Directly Observable Element, as

\[
P_{2,2}(k/k) = \frac{\|\tilde{x}_{2p}(k/k-1)\|^2 + \|\tilde{x}_1(k/k-1)\|^2}{\text{cov}[\tilde{x}_{2p}(k/k-1), \tilde{x}_1(k/k-1)]}
= \|\tilde{x}_{2p}(k/k-1)\|^2 + \|\tilde{x}_1(k/k-1)\|^2
- K_2^2(k)\|\tilde{y}(k)\|^2 \quad (37)
\]
which again is what also can be found from Eq. (13).

**Conclusion**

It is shown how the Kalman filter gain and covariance update equations can be illustrated geometrically. Though restricted to low dimensional systems the illustrations give an intuitive understanding of the filtering process. The covariance update equation illustrated can be considered a "natural" form in the sense that it directly relates to the error vectors involved.

**References**


---

**Thomas R. Kronhamn**

was born on January 30, 1943, in Gothenburg, Sweden. He received the M.Sc. degree in engineering physics from the Chalmers University of Technology in 1965. From 1966 to 1970, he was employed at Volvo Flygmotors in Trollhättan, Sweden, and participated in the development of advanced military turbofan engine control systems. Since 1970, he has been with Ericsson Radio Systems, Malmö, Sweden, where he has been engaged in the development of radar tracking systems for both aircraft and ground based applications. During the academic year 1983/84, he was a visiting scientist at the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology, doing research in the multitarget tracking area. Thomas Kronhamn is currently a Staff Scientist at the Signal and Data Processing Department responsible for coordination and long-term development of data processing algorithms.